CDF note 10761

Top Quark Mass Measurement in the Lepton+Jets Using 8.7 fb^{-1} of data

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: February 3, 2012)

We report on a measurement of the top quark mass (M_{top}) in the Lepton+Jets using $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV from 8.7 fb $^{-1}$ of data collected with the CDF detector at the Fermilab Tevatron. A top quark mass (m_t^{reco}) is reconstructed for every event by minimizing a χ^2-like function to the overconstrained kinematics of the $t\bar{t}$ system. The dijet mass (m_{jj}) of the hadronically decaying W boson is used to constrain *in situ* the largest systematic on top quark mass measurements, the uncertain jet energy scale (Δ_{JES}) in the detector. The additional information of top quark mass have been included by using a reconstructed top quark mass from 2^{nd} best χ^2 fit ($m_t^{reco(2)}$). The values of m_t^{reco}, m_{jj}, and $m_t^{reco(2)}$ for lepton+jets candidate events are compared to three-dimensional probability density function derived by applying kernel density estimation to fully simulated MC events with different values of the top quark mass and Δ_{JES} in the detector. We measure $M_{top} = 172.85 \pm 0.71$ (stat.) ± 0.84 (syst.) GeV/c^2.

Preliminary Results of TMT using 8.7 *fb* $^{-1}$

I. INTRODUCTION

This note describes a measurement of the mass of the top quark using $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV with the CDF detector at the Tevatron using full data of Run II update previous measurement [1]. The mass of the top quark is of much interest to particle physicists, both because the top quark is the heaviest known fundamental particle, and also because a precise measurement of the top quark mass helps constrain the mass of the Higgs boson. Top quarks are produced predominantly in pairs at the Tevatron, and in the Standard Model decay nearly 100% of the time to a W boson and a b quark. The topology of a $t\bar{t}$ event is determined by the decay of the two W bosons, as each W boson can decay to a lepton-neutrino pair (lν) or to a pair of quarks (qq'). We look for events consistent with $t\bar{t}$ production and decay involving one lν and the other qq' (we do not consider events with taus). Events in this final state, Lepton+Jets $t\bar{t}$ candidates, are consistent with which one W boson decays hadronically and the other W boson decays leptonically. The CDF detector is described in detail in [2].

Our measurement is a template-based measurement, meaning that we compare quantities in data with distributions from simulated MC events to find the most likely parent top quark mass distribution. In this measurement, we use three variables, two variables ($m_t^{\rm reco}$, $m_t^{\rm reco(2)}$) that is strongly correlated with the true top quark mass ($\rm M_{top}$) and the other variable ($m_{\rm jj}$) that is sensitive to shifts in the jet energy scale ($\Delta_{\rm JES}$) in the detector. The value of $m_t^{\rm reco}$ in each event is derived from a χ^2 minimization that uses knowledge of the overconstrained kinematics of the $t\bar{t}$ system [1, 3–5]. Because $m_t^{\rm reco}$ do not bring 100 % of the $\rm M_{top}$ information, we include another reconstructed top quark mass from 2^{nd} best combination of jets-to-parton from kinematic fit by choosing the 2^{nd} smallest χ^2 combination. The dijet mass ($m_{\rm jj}$) that we use in each event is chosen such that it often comes from the decay of the W resonance, and is sensitive to possible miscalibration of JES in the CDF detector.

Monte Carlo samples generated with 76 different $\rm M_{top}$ are run through a full CDF detector simulation assuming 29 possible shifts in $\Delta_{\rm JES}$. The values of observables in data are compared to each point in the MC grid using a non-parametric approach based on Kernel Density Estimation (KDE) [6]. Local Polynomial Smoothing [7] is used to smooth out these points and calculate the probability densities at any arbitrary value of $\rm M_{top}$ and $\Delta_{\rm JES}$. An unbinned likelihood fit is used to measure $\rm M_{top}$ and profile out $\Delta_{\rm JES}$.

II. EVENT SELECTION

At the trigger level, Lepton+Jets candidate events are selected by requiring a high-E_T electron (or high-p_T muon). In addition, large $\not\!\!E_T$ + two jets requirement is used to increase muon acceptance. Offline, the events are required to have a single energetic lepton (electron or muon), large $\not\!\!E_T$ due to the escaping neutrino from the leptonic W decay, and at least four jets in the final state. Electron candidates are identified as a high-momentum track in the tracking system matched to an electromagnetic cluster reconstructed in the calorimeters with $E_T > 20$ GeV. We also require that energy shared by the towers surrounding the cluster is low. Muon candidates are reconstructed as high-momentum tracks with $p_T > 20$ GeV/c matching hits in the muon chambers. Energy deposited in the calorimeter is required to be consistent with a minimum ionizing particle. The $\not\!\!E_T$ is required to be greater than 20 GeV. For the further rejection of QCD multijet background, we require the value of H_T greater than 250 GeV and minimum ϕ angle between $\not\!\!E_T$ and jets to be greater than 0.5 for zero and one b-tagged samples.

Jets are reconstructed with the JETCLU [8] cone algorithm using a cone radius of $R \equiv \sqrt{\eta^2 + \phi^2} = 0.4$. To improve the statistical power of the method, we divided sample into five subsamples, depending on the number of jets and the number of b-tagged jets. The SECVTX [9] algorithm uses the transverse decay length of tracks inside jets to tag jets as coming from b quarks. In Lepton+Jets events with exactly zero tag (0-tag), we require exactly four jets with $E_T > 20$ GeV/c^2 (tight jet). For events with one or more than one tag, which have more statistical power and less background, we loosen these cuts , and allow events with more than four jets. We also loosen the cut on the 4th jet to $E_T > 12$ GeV/c^2 to increase the number of such events. We separate samples based on whether we have exactly four tight jets (Tight) or not (Loose) and number of b-tagging. We then have five categories such as 0-tag (exactly zero b-tag and four tight jets), 1-tagT (exactly one b-tag and four tight jets), 1-tagL (exactly one b-tag and three or more than four tight jets), 2-tagT (more than one b-tag and exactly four tight jets), and 2-tagL (more than one b-tag and three or more than four tight jets).

For Lepton+Jets events, we make a cut on the χ^2 out of the kinematic fitter described in Section IV, requiring it to be less than 9.0 for tagged (1-tag and 2-tag) events and 3.0 for 0-tag events. In order to properly normalize our probability density functions, we define hard boundaries on the values of the observables. Events with values of an observable falling outside the boundary are rejected. Event selection and categorization of tight and loose samples are summarized in Tables I and II respectively.

TABLE I: Event selection in the Lepton+Jets channel.

	0-tag	1-tag	2-tag
b-tags	$= 0$	$= 1$	> 1
Leading 3 jets E_T (GeV)	> 20	> 20	> 20
Missing E_T (GeV)	> 20	> 20	> 20
4th jet E_T (GeV)	> 20	> 12	> 12
Extra jets E_T (GeV)	< 20	Any	Any
H_T (GeV)	> 250	> 250	Any
QCD veto	yes	yes	no
χ^2	< 3	< 9	< 9
$m_t^{\rm reco}$ boundary cut (GeV/c^2)	$100 < m_t^{\rm reco} < 350$	$100 < m_t^{\rm reco} < 350$	$100 < m_t^{\rm reco} < 350$
$m_{\rm jj}$ boundary cut (GeV/c^2)	$60 < m_{\rm jj} < 110$	$50 < m_{\rm jj} < 120$	$50 < m_{\rm jj} < 125$
$m_t^{\rm reco(2)}$ boundary cut (GeV/c^2)	$100 < m_t^{\rm reco(2)} < 350$	$100 < m_t^{\rm reco(2)} < 350$	$100 < m_t^{\rm reco(2)} < 350$

TABLE II: Event categories based on jet requirement.

	Loose	Tight
Leading 3 jets E_T (GeV)	> 20	> 20
η of leading 3 jets	< 2.0	< 2.0
4th jet E_T (GeV)	> 12	> 20
η of leading 3 jets	< 2.4	< 2.0
Extra jets E_T (GeV)	Any	< 20
number of tight jets	not 4	$= 4$

III. JET ENERGY SCALE

We describe in this section the *a priori* determination of the jet energy scale uncertainty by CDF that is used later in this analysis. More information on JES, calibration and uncertainty can be found in [10]. There are many sources of uncertainties related to jet energy scale at CDF:

- Relative response of the calorimeters as a function of pseudorapidity.
- Single particle response linearity in the calorimeters.
- Fragmentation of jets.
- Modeling of the underlying event energy.
- Amount of energy deposited out of the jet cone.

The uncertainty on each source is evaluated separately as a function of the jet p_T (and η for the first uncertainty in the list above). Their contributions are shown in Fig. 1 for the region $0.2 < \eta < 0$ 6. The black lines show the sum in quadrature (σ_c) of all contributions. This $\pm 1\sigma_c$ total uncertainty is taken as a unit of jet energy scale miscalibration (Δ_{JES}) in this analysis.

In addition to standard jet energy corrections, we employ neural network correction to use not only energy provided by the calorimeter but also momentum provided by the tracker that have been used Higgs search analysis at CDF [11]. Jet energy of b-quark is corrected separately for b-tagged events and non b-tagged events with parton level matching using PYTHIA signal MC. We also correct W daughter quark in the similar manner. This technique improve the resolution of jet energy resulting better precision of $\rm M_{top}$ measurement.

IV. LEPTON+JETS TOP MASS RECONSTRUCTION

The value of the reconstructed mass in each Lepton+Jets event ($m_t^{\rm reco}$) is determined by minimizing a χ^2 describing the overconstrained kinematics of the $t\bar{t}$ system. The reconstructed mass is a number that distills all the kinematic information in each event into one variable that is a good estimator for the true top quark mass. The kinematic fitter uses knowledge of the lepton and jet four-vectors, b-tagging information and the measured $\not\!\!E_T$. The invariant masses

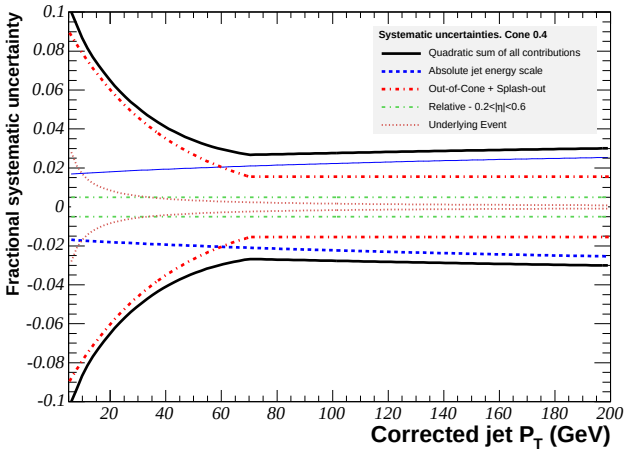

FIG. 1: Jet energy scale uncertainty as a function of the corrected jet p_T for the underlying event (dotted red), relative response (dashed green), out-of-cone energy (dashed red) and absolute response (dashed blue). The contribution of all sources are added in quadrature (full black) to form the total $\Delta_{\rm JES}$ systematic σ_c.

of the lepton-neutrino pair and the dijet mass from the hadronic W decay are constrained to be near the well-known W mass, and the two top quark masses per event are constrained to be equal within the narrow top width. The χ^2,

$$\begin{aligned}\chi^2 &= \Sigma_{i=\ell,4jets}\frac{(p_T^{i,fit}-p_T^{i,meas})^2}{\sigma_i^2}+\Sigma_{j=x,y}\frac{(U_j^{fit}-U_j^{meas})^2}{\sigma_j^2}\\ &+ \frac{(M_{jj}-M_W)^2}{\Gamma_W^2}+\frac{(M_{\ell\nu}-M_W)^2}{\Gamma_W^2}+\frac{(M_{bjj}-m_t^{\rm reco})^2}{\Gamma_t^2}+\frac{(M_{b\ell\nu}-m_t^{\rm reco})^2}{\Gamma_t^2}\end{aligned} \quad (1)$$

is minimized for every jet-parton assignment consistent with b-tagging. The first sum constrains the p_T of the jets and lepton, within their uncertainties, to remain close to their measured values. The second term constrains the unclustered energy in the event to remain near its measured value, providing a handle on the neutrino 4-vector. The W boson has a small width, and the two W mass terms provide the most powerful constraints in the fit. The last two terms in the χ^2 constrain the three-body invariant masses of each top decay chain to remain close to a single top quark mass, $m_t^{\rm reco}$. The single jet-parton assignment with the lowest χ^2 that is consistent with b-tagging gives the value of $m_t^{\rm reco}$ for the event. Events where the lowest $\chi^2 > 9.0$ are rejected.

Although $\rm m_t^{reco}$ carry a lot of information about $\rm M_{top}$, it is still possible to be carried $\rm M_{top}$ information by another combination of jet-parton assignment. Therefore, we add another $\rm M_{top}$ from the 2^{nd} smallest χ^2 combination which is $m_t^{\rm reco(2)}$. We use this value as 3^{rd} observable in the 3d KDE machinery.

V. DIJET MASS

The value of $m_{\rm jj}$ in each Lepton+Jets event can have an ambiguity due to not knowing which two jets came from a hadronic W decay. In 2-tag events, the value is chosen as the invariant mass of the two non-tagged jets in the leading 4 jets. In single-tag events, there are 3 dijet masses that can be formed from the 3 non-tagged jets among the 4 jets in the event. We chose the single dijet mass that is closest to the well known W mass.

VI. BACKGROUNDS

An *a priori* estimate for the Lepton+Jets background composition is used to derive background shapes for $m_t^{\rm reco}$, $m_t^{\rm reco(2)}$, and $m_{\rm jj}$. ALPGEN [12] combined with PYTHIA [13] is used to model W+jets. Contributions include W$b\bar{b}$, W$c\bar{c}$, Wc and W+light flavor (LF) jets. Anti-electron samples are used to model the QCD background. The relative fractions of the different W+jets samples are determined in MC, but the absolute normalization is derived from the

TABLE III: Expected numbers of background and signal events and observed events after event selection, χ^2 and boundary cuts for each category.

CDF II Preliminary 8.7 fb^{-1}

	0-tag	1-tagL	1-tagT	2-tagL	2-tagT
W$b\bar{b}$	37.6±15.9	54.4±22.6	34.0±14.3	8.5±3.6	6.1±2.6
W$c\bar{c}$	117.8±46.2	35.7±13.6	22.3±9.0	1.4±0.7	1.2±0.5
Wc	54.2±25.1	19.1±10.0	10.4±5.1	0.8±0.3	0.5±0.2
W+light jets	493.6±111.5	60.5±13.5	35.4±9.0	0.9±0.3	0.6±0.2
Z+jets	52.3±4.4	8.9±1.1	5.9±0.7	0.8±0.1	0.5±0.1
single top	4.9±0.5	10.5±0.9	6.8±0.6	2.2±0.3	1.7±0.2
Diboson	60.3±5.6	11.1±1.4	8.5±1.1	1.0±0.2	0.8±0.1
QCD	143.0±114.4	34.5±12.6	20.7±16.6	4.4±2.5	2.5±2.4
Total	963.5±229.3	234.7±61.1	144.0±40.9	19.9±5.5	13.8±4.2
$t\bar{t}$	644.8±86.3	695.0±86.7	867.3±107.6	192.3±29.7	303.7±46.6
Expected Events	1608.4±245.0	929.8±106.1	1011.3±115.1	212.2±30.2	317.6±46.8
Observed Events	1627	882	997	208	275

data. The MC are combined using their relative cross sections and acceptances, and we remove events overlapping in phase space and flavor across different samples. MC and theoretical cross-sections are used to model the single-top and diboson backgrounds. The expected number of background from different sources is shown in Table III. The backgrounds are assumed to have no M_{top} dependence, but all MC-based backgrounds are allowed to have Δ_{JES} dependence.

VII. KERNEL DENSITY ESTIMATES

Probability density functions for $m_t^{\rm reco}$-$m_{\rm jj}$-$m_t^{\rm reco(2)}$ at every point in the $M_{top} - \Delta_{JES}$ grid and for backgrounds are derived using a Kernel Density Estimate (KDE) approach [6]. KDE is a non-parametric method for forming density estimates that can easily be generalized to more than one dimension, making it useful for this analysis, which has two observables per event. The probability for an event with observable (x) is given by the linear sum of contributions from all entries in the MC:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x - x_i}{h}). \tag{2}$$

In the above equation, $\hat{f}(x)$ is the probability to observe x given some MC sample with known mass and JES (or the background). The MC has n entries, with observables x_i. The kernel function K is a normalized function that adds less probability to a measurement at x as its distance from x_i increases. The smoothing parameter h (sometimes called the bandwidth) is a number that determines the width of the kernel. Larger values of h smooth out the contribution to the density estimate and give more weight at x farther from x_i. Smaller values of h provide less bias to the density estimate, but are more sensitive to statistical fluctuations. We use the Epanechnikov kernel, defined as:

$$K(t) = \frac{3}{4}(1 - t^2) \text{ for } |t| < 1 \text{ and } K(t) = 0 \text{ otherwise,} \tag{3}$$

so that only events with $|x - x_i| < h$ contribute to $\hat{f}(x)$. We use an adaptive KDE method in which the value of h is replaced by h_i in that the amount of smoothing around x_i depends on the value of $\hat{f}(x_i)$. In the peak of the distributions, where statistics are high, we use small values of h_i to capture as much shape information as possible. In the tails of the distribution, where there are few events and the density estimates are sensitive to statistical fluctuations, a larger value of h_i is used. The overall scale of h is set by the number of entries in the MC sample (larger smoothing is used when fewer events are available), and by the RMS of the distribution (larger smoothing is used for wider distributions). We extend KDE to three dimensions by multiplying the three kernels together.

$$\hat{f}(x,y,z) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h_{x,i}h_{y,i}h_{z,i}} \left[K(\frac{x - x_i}{h_{x,i}}) \times K(\frac{y - y_i}{h_{y,i}}) \times K(\frac{z - z_i}{h_{z,i}}) \right]. \tag{4}$$

VIII. LIKELIHOOD FIT

The values of $m_t^{\rm reco}$-$m_{\rm jj}$-$m_t^{\rm reco(2)}$ and $m_t^{\rm NWA}$-$m_{\rm T2}$ observed in data are compared to points in $\rm M_{top} - \Delta_{JES}$ space. An maximum likelihood fit is performed to maximize the likelihood with respect to the expected number of signal (n_s) and background events (n_b) in each of the four subsamples. A Gaussian constraint on the expected number of background events is applied to each of the subsamples. The likelihood for subsample k with N events is given by:

$$\mathcal{L}_k = \exp\left(-\frac{(n_b - n_b^0)^2}{2\sigma_{n_b}^2}\right) \times \prod_{i=1}^{N} \frac{n_s P_{sig}(m_t^{\rm reco}, m_{\rm jj}, m_t^{\rm reco(2)}; M_{top}, \Delta_{JES}) + n_b P_{bg}(m_t^{\rm reco}, m_{\rm jj}, m_t^{\rm reco(2)}; \Delta_{JES})}{n_s + n_b} \quad (5)$$

The overall likelihood is a product over the five individual subsample likelihoods, with a Gaussian constraint on $\Delta_{\rm JES}$, constraining it to the nominal $0 \pm 1\ \sigma_c$:

$$\mathcal{L} = \exp\left(-\frac{\Delta_{\rm JES}^2}{2{\sigma_c}^2}\right) \times \mathcal{L}_{\text{0-tag}} \times \mathcal{L}_{\text{1-tagL}} \times \mathcal{L}_{\text{0-tagT}} \times \mathcal{L}_{\text{2-tagL}} \times \mathcal{L}_{\text{2-tagT}}. \quad (6)$$

The above gives values of $-\ln\mathcal{L}$ only for points in the $\rm M_{top} - \Delta_{JES}$ grid, and not as a continuous function. To obtain density estimates for an arbitrary point in the $\rm M_{top} - \Delta_{JES}$ grid, we use local polynomial smoothing [7] on a per-event basis. The value of the density estimate is obtained for an event at the available points, and a quadratic fit is performed in $\rm M_{top} - \Delta_{JES}$ space, where the values of $\rm M_{top}$ and $\Delta_{\rm JES}$ far away from the point being estimated are deweighted. This allows for a smooth likelihood that can be minimized. The measured uncertainty on $\rm M_{top}$ comes from the largest possible shift in $\rm M_{top}$ on the $\Delta\ln\mathcal{L} = 0.5$ contour.

IX. METHOD CHECK

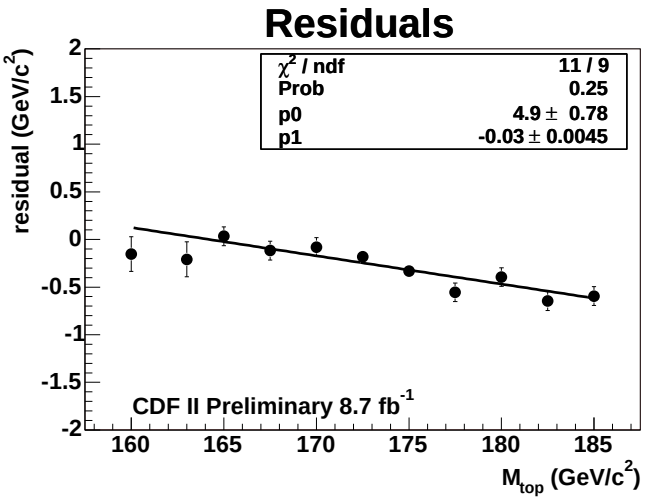

FIG. 2: Residual mass shift as a function of input mass from pseudoexperiments before corrections.

We test our machinery by running pseudoexperiments with varying values of $\rm M_{top}$ between 160 and 182.5 GeV/c^2 and varying values of $\Delta_{\rm JES}$ between -1.0 and 1.0 σ_c. Figure 2 shows the $\rm M_{top}$ residuals as a function of true top quark mass and it is biased. However, it can be well fitted by linear function so, we correct bias to be matched with zero. This correction affect pull widths to be off from unity, so we need to inflate our measured statistical uncertainties by 2.9%. Figure 3 shows $\rm M_{top}$ residuals after correction for each measurement. Pull width with errrors inflation are shown in Fig. 4. For the all of measurement including systematic study, we apply these corrections.

X. RESULTS

The likelihood procedure when applied to the data yields $M_{top} = 172.85 \pm 0.71$ GeV/c^2. The log-likelihood contours is shown in Figures 5. As shown in Fig. 6, 17 % of pseudoexperiments have a smaller error than the value measured in the data fit. Figure 7 shows the measured distributions of the observables used for M_{top} measurement overlaid with density estimates using $t\bar{t}$ signal events with M_{top} =173 GeV/c^2 and Δ_{JES} =0.0 with the full background model.

The fit returns $\Delta_{JES} = 0.11 \pm 0.15\ \sigma_c$.

We run a fit separately without the JES and background constraints and measure $M_{top} = 172.84 \pm 0.71$ GeV/c^2 and $M_{top} = 172.77 \pm 0.70$ GeV/c^2, showing that these priors do not significantly affect our result.

XI. SYSTEMATIC UNCERTAINTIES

We examine a variety of effects that could systematically shift our measurement. As a single nuisance parameter, the JES that we measure does not fully capture the complexities of possible jet energy scale uncertainties, particularly those with different η and p_T dependence. Fitting for the global JES removes most of these effects, but not all of them. We apply variations within uncertainties to different JES calibrations for the separate known effects in both signal and background pseudodata and measure resulting shifts in M_{top} from pseudoexperiments, giving a residual JES uncertainty. For the dilepton-only measurement, which has no *in situ* calibration, these systematics dominate. We also vary the energy of b jets, which have different fragmentation than light quarks jets, as well as semi-leptonic decays and different color flow, resulting in a b-JES systematic. Effects due to uncertain modeling of radiation including initial-state radiation (ISR) and final-state radiation (FSR) are studied by extrapolating uncertainties in the p_T of Drell-Yan events to the $t\bar{t}$ mass region, resulting in a radiation systematics. Comparing pseudoexperiments generated with HERWIG [14] and PYTHIA gives an estimate of the generator systematic. A systematic on different parton distribution functions is obtained by varying the independent eigenvector of the CTEQ6M set, comparing parton distribution functions with different values of Λ_{QCD}, and comparing CTEQ5L with MRST72. We also test the effect of reweighting MC to increase the fraction of $t\bar{t}$ events initiated by gg (vs qq) from the 6% in the leading order MC to 20%. Systematics due to lepton energy scales are estimated by propagating 1% shifts on electron and muon energies scales. Background composition systematics are obtained by varying the fraction of the different types of backgrounds in pseudoexperiments. Varying the uncertain Q^2 of background events results in a background shape systematic. The next leading order effect is evaluated by using MC@NLO [15] MC compared with HERWIG MC. B-tagging efficiency gives the systematic effect by varying the efficiency within its uncertainty as a function of jet p_T. It has been suggested that color reconnection effects could cause a bias in the top quark mass measurement [16]. We test this effect by generating MCs with and without CR and take the difference as systematics.

The total systematic error is 0.8 GeV/c^2. The systematics are summarized in Table IV.

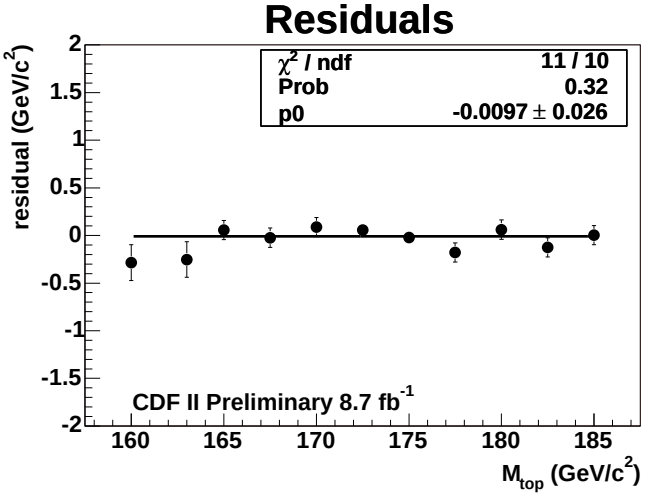

FIG. 3: Residual mass shift as a function of input mass from pseudoexperiments after corrections.

TABLE IV: Summary of systematics. All numbers have units of GeV/c^2.

CDF II Preliminary 8.7 fb^{-1}

Systematic	GeV/c^2
Residual JES	0.52
Generator	0.56
Next Leading Order	0.09
PDFs	0.08
b jet energy	0.10
b tagging efficiency	0.03
Background shape	0.20
gg fraction	0.03
Radiation	0.06
MC statistics	0.05
Lepton energy	0.03
MHI	0.07
Color Reconnection	0.21
Total systematic	0.84

XII. CONCLUSIONS

We present a simultaneous measurement of the mass of the top quark in the Lepton+Jets channel using a template-based technique with an *in situ* JES calibration. Using 3d templates derived from Kernel Density Estimation and 8.7 fb^{-1} of data collected by the Tevatron, we measure

$$\mathrm{M_{top}} = 172.85 \pm 0.71\ (\mathrm{stat.}) \pm 0.84\ (\mathrm{syst.})\ \mathrm{GeV}/c^2 = 172.85 \pm 1.10\ \mathrm{GeV}/c^2.$$

Acknowledgments

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of

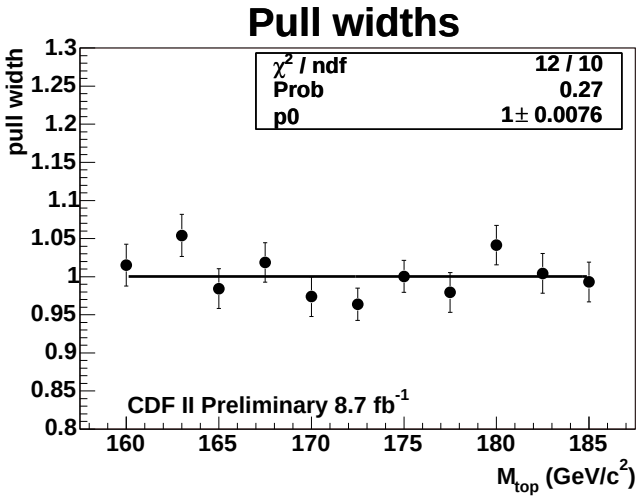

FIG. 4: Pull widths as a function of input mass from pseudoexperiments. The pull widths already have inflation of statistical uncertainties by 2.9%.

China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; and the Academy of Finland.

[1] T. Aaltonen, el al., Phys. Rev. D **83**, 111101 (2011);
[2] F. Abe, et al., Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988);
[3] T. Aaltonen, el al., Phys. Rev. D **79**, 092005 (2009);
[4] A. Abulenci et al., Phys. Rev. D 73, 032003 (2006);
[5] A. Abulenci et al., Phys. Rev. Lett. 96, 022004 (2006);
[6] K. S. Cranmer, Comput. Phys. Commun. **136**, 198 (2001);
[7] C. Loader, *Local Regression and Likelihood*, New York, USA, (Springer, 1999);
[8] F. Abe, et al., Phys. Rev. D 45, 1448 (1992);
[9] T. Affolder, et al., Phys. Rev. D 64, 032002 (2001);
[10] A. Bhatti, et al., Nucl. Instrum. Methods Phys. Res. A **566**, 375 (2006);
[11] T. Aaltonen, et al., arXiv:1107.3026;
[12] M. L. Mangano, et al., J. High Energy Phys. **0307**, 001 (2003);
[13] T. Sjostrand, S. Mrenna, and P. Skands, J. High Energy Phys. **0605**, 026 (2006);
[14] G. Corcella, et al., J. High Energy Phys. **0101**, 010 (2001);
[15] S. Frixione and B.R. Webber, J. High Energy Phys. **0206**, 029 (2002);
[16] P. Z. Skands, D. Wicke, Eur. Phys. J. C **52**, 133 (2007);

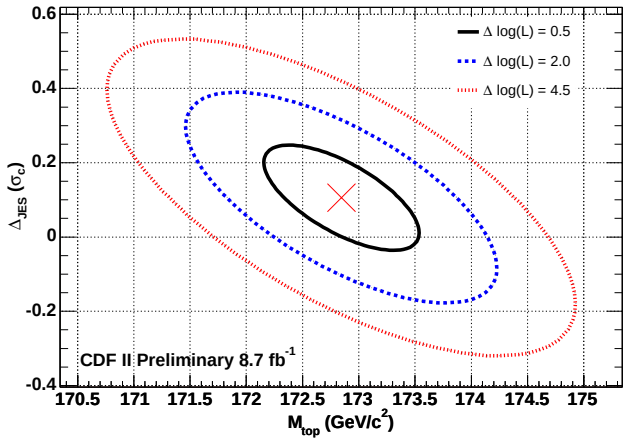

FIG. 5: Negative log-likelihood contour of data fit.

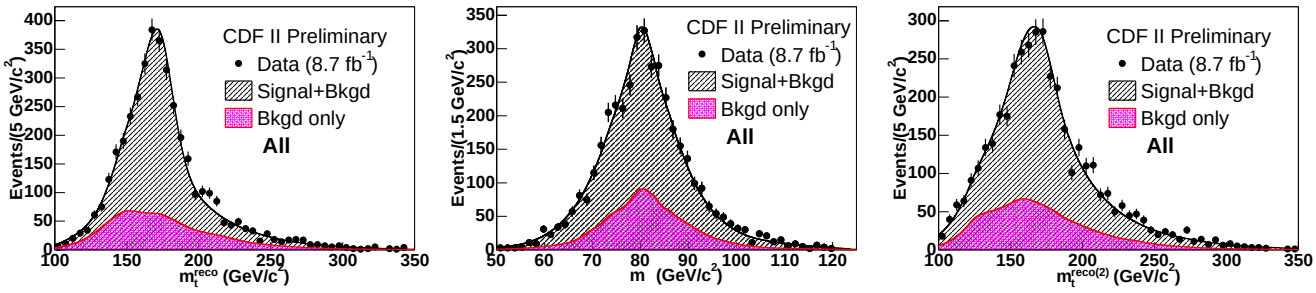

FIG. 7: Distributions of $m_t^{\rm reco}$, m_{jj}, and $m_t^{\rm reco(2)}$ used to extract $M_{\rm top}$ for events combining all categories. The data is overlaid with the predictions from the KDE probability distributions assuming $M_{\rm top}$ =173 GeV/c^2 and $\Delta_{\rm JES}$ =0.0.

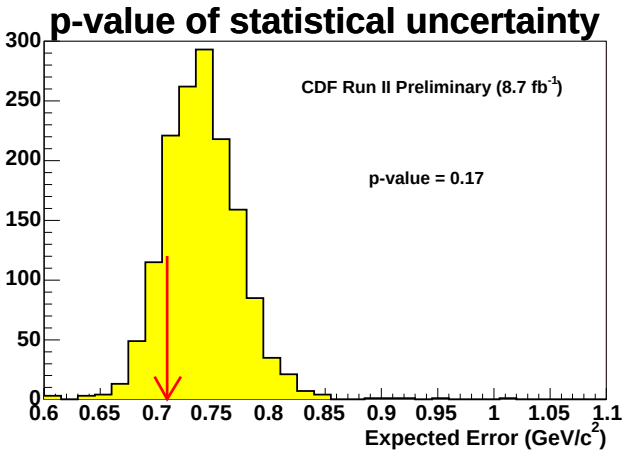

FIG. 6: Reported error from pseudoexperiments with the observed number of events.